PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12P 19/34, C12Q 1/68, C12M 1/34

A1

(11) International Publication Number: **WO 00/70073**

(43) International Publication Date: 23 November 2000 (23.11.00)

(21) International Application Number: PCT/US00/13677

(22) International Filing Date: 18 May 2000 (18.05.00)

(30) Priority Data:
60/134,827 19 May 1999 (19.05.99) US

(71) Applicant: CORNELL RESEARCH FOUNDATION, INC. [US/US]; Suite 105, 20 Thornwood Drive, Ithaca, NY 14850 (US).

(72) Inventors: KORLACH, Jonas; 2108 Hasbrouck Apt., Ithaca, NY 14850 (US). WEBB, Watt, W.; 9 Parkway Place, Ithaca, NY 14850 (US). LEVENE, Michael; 700 Warren Road, 11-3D, Ithaca, NY 14850 (US). TURNER, Stephen; 17 Sheraton Drive, Ithaca, NY 14850 (US). CRAIGHEAD, Harold, G.; 21 Fairway Drive, Ithaca, NY 14850 (US). FOQUET, Mathieu; 21 Fairway Drive, Ithaca, NY 14850 (US).

(74) Agents: GOLDMAN, Michael, L. et al.; Nixon Peabody LLP, Clinton Square, P.O. Box 1051, Rochester, NY 14603 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: METHOD FOR SEQUENCING NUCLEIC ACID MOLECULES

(57) Abstract

The present invention is directed to a method of sequencing a target nucleic acid molecule having a plurality of bases. In its principle, the *temporal* order of base additions during the polymerization reaction is measured on a molecule of nucleic acid, i.e. the activity of a nucleic acid polymerizing enzyme on the templage nucleic acid molecule to be sequenced is followed in real time. The sequence is deduced by identifying which base is being incorporated into the growing complementary strand of the target nucleic acid by the catalytic activity of the nucleic acid polymerizing enzyme at each step in the sequence of base additions. A polymerase on the target nucleic acid molecule complex is provided in a position suitable to move along the target nucleic acid molecule and extend the oligonucleotide primer at an active site. A plurality of labelled types of nucleotide analogs are provided proximate to the active site, with each distinguishable type of nucleotide analog being complementary to a different nucleotide in the target nucleic acid sequence. The growing nucleic acid strand is extended by using the polymerase to add a nucleotide analog to the nucleic acid strand at the active site, where the nucleotide analog being added is complementary to the nucleotide of the target nucleic acid at the active site. The nucleotide analog added to the oligonucleotide primer as a result of the polymerizing step is identified. The steps of providing labelled nucleotide analogs, polymerizing the growing nucleic acid strand, and identifying the added nucleotide analog are repeated so that the nucleic acid strand is further extended and the sequence of the target nucleic acid is determined.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

METHOD FOR SEQUENCING NUCLEIC ACID MOLECULES

This application claims benefit of U.S. Provisional Patent Application Serial No. 60/134,827, filed May 19, 1999.

This invention was made with funds provided by the U.S. Government under National Science Foundation Grant No. BIR8800278, National Institutes of Health Grant No. P412RR04224-11, and Department of Energy Grant No. 066898-0003891. The U.S. Government may have certain rights in this invention.

FIELD OF THE INVENTION

The present invention relates to a method for determining the sequence of nucleic acid molecules.

BACKGROUND OF THE INVENTION

The goal to elucidate the entire human genome has created an interest in technologies for rapid DNA sequencing, both for small and large scale applications. Important parameters are sequencing speed, length of sequence that can be read during a single sequencing run, and amount of nucleic acid template required. These research challenges suggest aiming to sequence the genetic information of single cells without prior amplification, and without the prior need to clone the genetic material into sequencing vectors. Large scale genome projects are currently too expensive to realistically be carried out for a large number of organisms or patients. Furthermore, as knowledge of the genetic basis for human diseases increases, there will be an ever-increasing need for accurate, high-throughput DNA sequencing that is affordable for clinical applications. Practical methods for determining the base pair sequences of single molecules of nucleic acids, preferably with high speed and long read lengths, would provide the necessary measurement capability.

Two traditional techniques for sequencing DNA are the dideoxy termination method of Sanger (Sanger et al., Proc. Natl. Acad. Sci. U.S.A. 74: 563-5467 (1977)) and the Maxam-Gilbert chemical degradation method (Maxam and

Gilbert, Proc. Natl. Acad. Sci. U.S.A. 74: 560-564 (1977)). Both methods deliver four samples with each sample containing a family of DNA strands in which all strands terminate in the same nucleotide. Ultrathin slab gel electrophoresis, or more recently capillary array electrophoresis is used to resolve the different length strands and to determine the nucleotide sequence, either by differentially tagging the strands of each sample before electrophoresis to indicate the terminal nucleotide, or by running the samples in different lanes of the gel or in different capillaries. Both the Sanger and the Maxam-Gilbert methods are labor- and time-intensive, and require extensive pretreatment of the DNA source. Attempts have been made to use mass spectroscopy to replace the time-intensive electrophoresis step. For review of existing sequencing technologies, see Cheng "High-Speed DNA-Sequence Analysis," Prog. Biochem. Biophys. 22: 223-227 (1995).

Related methods using dyes or fluorescent labels associated with the terminal nucleotide have been developed, where sequence determination is also made by gel electrophoresis and automated fluorescent detectors. For example, the Sanger-extension method has recently been modified for use in an automated micro-sequencing system which requires only sub-microliter volumes of reagents and dye-labelled dideoxyribonucleotide triphosphates. In U.S. Patent No. 5,846,727 to Soper et al., fluorescence detection is performed on-chip with one single-mode optical fiber carrying the excitation light to the capillary channel, and a second single-mode optical fiber collecting the fluorescent photons. Sequence reads are estimated in the range of 400-500 bases which is not a significant improvement over the amount of sequence information obtained with traditional Sanger or Maxam-Gilbert methods. Furthermore, the Soper method requires PCR amplification of template DNA, and purification and gel electrophoresis of the oligonucleotide sequencing 'ladders,' prior to initiation of the separation reaction. These systems all require significant quantities of target DNA. Even the method described in U.S. Patent No. 5,302,509 to Cheeseman, which does not use gel electrophoresis for sequence determination, requires at least a million DNA molecules.

In a recent improvement of a sequencing-by-synthesis methodology originally devised ten years ago, DNA sequences are being deduced by measuring pyrophosphate release upon testing DNA/polymerase complexes with each

deoxyribonucleotide triphosphate (dNTP) separately and sequentially. See Ronaghi et al., "A Sequencing Method Based on Real-Time Pyrophosphate," Science 281: 363-365 (1998) and Hyman, "A New Method of Sequencing DNA," Anal. Biochem. 174: 423-436 (1988). While using native nucleotides, the method requires synchronization of polymerases on the DNA strands which greatly restricts sequence read lengths. Only about 40 nucleotide reads were achieved, and it is not expected that the detection method can approach single molecule sensitivity due to limited quantum efficiency of light production by luciferase in the procedure presented by Ronaghi et al., "A Sequencing Method Based on Real-Time Pyrophosphate," Science 281: 363-365 (1998). Furthermore, the overall sequencing speed is limited by the necessary washing steps, subsequent chemical steps in order to identify pyrophosphate presence, and by the inherent time required to test each base pair to be sequenced with all the four bases sequentially. Also, difficulties in accurately determining homonucleotide stretches in the sequences were recognized.

Previous attempts for *single molecule* sequencing (generally unsuccessful but seminal) have utilized exonucleases to sequentially release individual fluorescently labelled bases as a second step after DNA polymerase has formed a complete complementary strand. See Goodwin et al., "Application of Single Molecule Detection to DNA Sequencing," Nucleos. Nucleot. 16: 543-550 (1997). It consists of synthesizing a DNA strand labelled with four different fluorescent dNTP analogs, subsequent degradation of the labelled strand by the action of an exonuclease, and detection of the individual released bases in a hydrodynamic flow detector. However, both polymerase and exonuclease have to show activity on a highly modified DNA strand, and the generation of a DNA strand substituted with four different fluorescent dNTP analogs has not yet been achieved. See Dapprich et al., "DNA Attachment to Optically Trapped Beads in Microstructures Monitored by Bead Displacement," Bioimaging 6: 25-32 (1998). Furthermore, little precise information is known about the relation between the degree of labeling of DNA and inhibition of exonuclease activity. See Dörre et al., "Techniques for Single Molecule Sequencing," Bioimaging 5: 139-152 (1997).

In a second approach utilizing exonucleases, *native* DNA is digested while it is being pulled through a thin liquid film in order to spatially separate the

cleaved nucleotides. See Dapprich et al., "DNA Attachment to Optically Trapped Beads in Microstructures Monitored by Bead Displacement," Bioimaging 6: 25-32 (1998). They then diffuse a short distance before becoming immobilized on a surface for detection. However, most exonucleases exhibit sequence- and structure-dependent cleavage rates, resulting in difficulties in data analysis and matching sets from partial sequences. In addition, ways to identify the bases on the detection surface still have to be developed or improved.

Regardless of the detection system, methods which utilize exonucleases have not been developed into methods that meet today's demand for rapid, high-throughput sequencing. In addition, most exonucleases have relatively slow turnover rates, and the proposed methods require extensive pretreatment, labeling and subsequent immobilization of the template DNA on the bead in the flowing stream of fluid, all of which make a realization into a simple high-throughput system more complicated.

Other, more direct approaches to DNA sequencing have been attempted, such as determining the *spatial* sequence of fixed and stretched DNA molecules by scanned atomic probe microscopy. Problems encountered with using these methods consist in the narrow spacing of the bases in the DNA molecule (only 0.34 nm) and their small physicochemical differences to be recognized by these methods. See Hansma et al., "Reproducible Imaging and Dissection of Plasmid DNA Under Liquid with the Atomic Force Microscope," Science 256: 1180-1184 (1992).

In a recent approach for microsequencing using polymerase, but not exonuclease, a set of identical single stranded DNA (ssDNA) molecules are linked to a substrate and the sequence is determined by repeating a series of reactions using fluorescently labelled dNTPs. U.S. Patent No. 5,302,509 to Cheeseman. However, this method requires that each base is added with a fluorescent label and 3'-dNTP blocking groups. After the base is added and detected, the fluorescent label and the blocking group are removed, and, then, the next base is added to the polymer.

Thus, the current sequencing methods either require both polymerase and exonuclease activity to deduce the sequence or rely on polymerase alone with additional steps of adding and removing 3'-blocked dNTPs. The human genome project has intensified the demand for rapid, small- and large-scale DNA sequencing

that will allow high throughput with minimal starting material. There also remains a need to provide a method for sequencing nucleic acid molecules that requires only polymerase activity, without the use of blocking substituents, resulting in greater simplicity, easier miniaturizability, and compatibility to parallel processing of a single-step technique.

The present invention is directed to meeting the needs and overcoming deficiencies in the art.

SUMMARY OF THE INVENTION

The present invention relates to a method of sequencing a target nucleic acid molecule having a plurality of nucleotide bases. This method involves providing a complex of a nucleic acid polymerizing enzyme and the target nucleic acid molecule oriented with respect to each other in a position suitable to add a nucleotide analog at an active site complementary to the target nucleic acid. A plurality of types of nucleotide analogs are provided proximate to the active site, wherein each type of nucleotide analog is complementary to a different nucleotide in the target nucleic acid sequence. A nucleotide analog is polymerized at an active site, wherein the nucleotide analog being added is complementary to the nucleotide of the target nucleic acid, leaving the added nucleotide analog ready for subsequent addition of nucleotide analogs. The nucleotide analog added at the active site as a result of the polymerizing step is identified. The steps of providing a plurality of nucleotide analogs, polymerizing, and identifying are repeated so that the sequence of the target nucleic acid is determined.

Another aspect of the present invention relates to an apparatus suitable for sequencing a target nucleic acid molecule. This apparatus includes a support as well as a nucleic acid polymerizing enzyme or oligonucleotide primer suitable to bind to a target nucleic acid molecule, where the polymerase or oligonucleotide primer is positioned on the support. A microstructure defines a confined region containing the support and the nucleic acid polymerizing enzyme or the oligonucleotide primer which is configured to permit labeled nucleotide analogs that are not positioned on the support to move rapidly through the confined region.

A further feature of the present invention involves an apparatus suitable for sequencing a target nucleic acid molecule. This apparatus includes a solid support and a nucleic acid polymerizing enzyme or oligonucleotide primer suitable to hybridize to a target nucleic acid molecule, where the nucleic acid polymerizing enzyme or oligonucleotide primer is positioned on the support. A housing defines a confined region containing the support and the nucleic acid polymerizing enzyme or the oligonucleotide primer. The housing is constructed to facilitate identification of labeled nucleotide analogs positioned on the support. Optical waveguides proximate to the confined region focus activating radiation on the confined region and collect radiation from the confined region.

Numerous advantages are achieved with the present invention. Sequencing can be carried out with small amounts of nucleic acid, with the capability of sequencing single nucleic acid template molecules which eliminates the need for amplification prior to initiation of sequencing. Long read lengths of sequence can be deduced in one run, eliminating the need for extensive computational methods to assemble a gap-free full length sequence of long template molecules (e.g., bacterial artificial chromosome (BAC) clones). For two operational modes of the present inventions, the read length of the sequence is limited by the length of template to be sequenced, or the processivity of the polymerase, respectively. By using the appropriate enzymatic systems, e.g. with accessory proteins to initiate the sequencing reaction at specific sites (e.g., origins of replication) on the double-stranded template nucleic acid, preparative steps necessary for conventional sequencing techniques, such as subcloning into sequencing vectors, can be eliminated.

In addition, the sequencing method of the present invention can be carried out using polymerase and no exonuclease. This results in greater simplicity, easier miniaturizability, and compatibility to parallel processing of a single-step technique.

In regard to the latter advantage, some polymerases exhibit higher processivity and catalytic speeds than exonucleases, with over 10,000 bases being added before dissociation of the enzyme for the case of T7 DNA polymerase (compared to 3,000 bases for λ exonuclease). In some cases, e.g., T7 DNA polymerase complexed with T7 helicase/primase, processivity values are even higher,

ranging into several 100,000s. The rates of DNA synthesis can be very high, measured *in vivo* of 1,000 bases/sec and *in vitro* of 750 bases/sec (in contrast to 12 bases/sec degraded by λ exonuclease *in vitro*). See Kelman et al., "Processivity of DNA Polymerases: Two Mechanisms, One Goal," Structure 6: 121-125 (1998); Carter et al., "The Role of Exonuclease and Beta Protein of Phage Lambda in Genetic Recombination. II. Substrate Specificity and the Mode of Action of Lambda Exonuclease," J. Biol. Chem. 246: 2502-2512 (1971); Tabor et al., "Escherichia coli Thioredoxin Confers Processivity on the DNA Polymerase Activity of the Gene 5 Protein of Bacteriophage T7," J. Biol. Chem. 262: 16212-16223 (1987); and Kovall et al., "Toroidal Structure of Lambda-Exonuclease" Science 277: 1824-1827 (1997), which are hereby incorporated by reference. An incorporation rate of 750 bases/sec is approximately 150 times faster than the sequencing speed of one of the fully automated ABI PRISM 3700 DNA sequencers by Perkin Elmer Corp., Foster City, California, proposed to be utilized in a shot-gun sequencing strategy for the human genome. See Venter et al., "Shotgun Sequencing of the Human Genome," Science 280: 1540-1542 (1998), which is hereby incorporated by reference.

The small size of the apparatus that can be used to carry out the sequencing method of the present invention is also highly advantageous. The confined region of the template/polymerase complex can be provided by the microstructure apparatus with the possibility of arrays enabling a highly parallel operational mode, with thousands of sequencing reactions carried out sequentially or simultaneously. This provides a fast and ultrasensitive tool for research application as well as in medical diagnostics.

BRIEF DESCRIPTION OF DRAWINGS

Figures 1A-C show 3 alternative embodiments for sequencing in accordance with the present invention.

Figures 2A-C are schematic drawings showing the succession of steps used to sequence nucleic acids in accordance with the present invention.

Figures 3A-C show plots of fluorescence signals vs. time during the succession of steps used to sequence the nucleic acid in accordance with the present invention. Figure 3C shows the sequence generated by these steps.

Figures 4A-D depict the structure and schematic drawings showing the succession of steps used to sequence the nucleic acid in accordance with the present invention in the case where fluorescent nucleotides carrying the label at the gamma phosphate position (here shown as a gamma-linked dNTP) are used.

Figure 5 shows the principle of discrimination of fluorophores by time-gated fluorescence decay time measurements, which can be used to suppress background signal in accordance with the present invention.

Figure 6A shows a system for sequencing in accordance with the present invention. Figure 6B is an enlargement of a portion of that system.

Figure 7A shows a system for sequencing in accordance with the present invention using electromagnetic field enhancement with metal tips. Figure 7B is an enlargement of a portion of that system.

Figure 8A shows a system for sequencing in accordance with the present invention using near field apertures. Figure 8B is an enlargement of a portion of that system.

Figure 9A shows a system for sequencing in accordance with the present invention using nanochannels. Figure 9B is an enlargement of a portion of that system.

Figures 10A-B show systems for supplying reagents to a nanofabricated confinement system in accordance with the present invention. In particular, Figure 10A is a schematic drawing which shows how reagents are provided and passed through the system. Figure 10B is similar but shows this system on a single chip with pads to connect the system to fluid reservoirs.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a method of sequencing a target nucleic acid molecule having a plurality of nucleotide bases. This method involves providing a complex of a nucleic acid polymerizing enzyme and the target nucleic

acid molecule oriented with respect to each other in a position suitable to add a nucleotide analog at an active site complementary to the target nucleic acid. A plurality of types of nucleotide analogs are provided proximate to the active site, wherein each type of nucleotide analog is complementary to a different nucleotide in the target nucleic acid sequence. A nucleotide analog is polymerized at an active site, wherein the nucleotide analog being added is complementary to the nucleotide of the target nucleic acid, leaving the added nucleotide analog ready for subsequent addition of nucleotide analogs. The nucleotide analog added at the active site as a result of the polymerizing step is identified. The steps of providing a plurality of nucleotide analogs, polymerizing, and identifying are repeated so that the sequence of the target nucleic acid is determined.

Another aspect of the present invention relates to an apparatus suitable for sequencing a target nucleic acid molecule. This apparatus includes a support as well as a nucleic acid polymerizing enzyme or oligonucleotide primer suitable to bind to a target nucleic acid molecule, where the polymerase or oligonucleotide primer is positioned on the support. A microstructure defines a confined region containing the support and the nucleic acid polymerizing enzyme or the oligonucleotide primer which is configured to permit labeled nucleotide analogs that are not positioned on the support to move rapidly through the confined region.

A further feature of the present invention involves an apparatus suitable for sequencing a target nucleic acid molecule. This apparatus includes a support and a nucleic acid polymerizing enzyme or oligonucleotide primer suitable to hybridize to a target nucleic acid molecule, where the nucleic acid polymerizing enzyme or oligonucleotide primer is positioned on the support. A housing defines a confined region containing the support and the nucleic acid polymerizing enzyme or the oligonucleotide primer. The housing is constructed to facilitate identification of labeled nucleotide analogs positioned on the support. Optical waveguides proximate to the confined region focus activating radiation on the confined region and collect radiation from the confined region.

The present invention is directed to a method of sequencing a target nucleic acid molecule having a plurality of bases. In its fundamental principle, the *temporal* order of base additions during the polymerization reaction is measured on a

single molecule of nucleic acid, i.e. the activity of a nucleic acid polymerizing enzyme, hereafter also referred to as polymerase, on the template nucleic acid molecule to be sequenced is followed in real time. The sequence is deduced by identifying which base is being incorporated into the growing complementary strand of the target nucleic acid by the catalytic activity of the nucleic acid polymerizing enzyme at each step in the sequence of base additions. In the preferred embodiment of the invention, recognition of the time sequence of base additions is achieved by detecting fluorescence from appropriately labelled nucleotide analogs as they are incorporated into the growing nucleic acid strand. Accuracy of base pairing is provided by the specificity of the enzyme, with error rates of false base pairing of 10^{-5} or less. For enzyme fidelity, see Johnson, "Conformational Coupling in DNA-Polymerase Fidelity," Ann. Rev. Biochem. 62:685-713 (1993) and Kunkel, "DNA-Replication Fidelity," J. Biol. Chem. 267:18251-18254 (1992), which are hereby incorporated by reference.

The invention applies equally to sequencing all types of nucleic acids (DNA, RNA, DNA/RNA hybrids etc.) using a number of polymerizing enzymes (DNA polymerases, RNA polymerases, reverse transcriptases, mixtures, etc.). Therefore, appropriate nucleotide analogs serving as substrate molecules for the nucleic acid polymerizing enzyme can consist of members of the groups of dNTPs, NTPs, modified dNTPs or NTPs, peptide nucleotides, modified peptide nucleotides, or modified phosphate-sugar backbone nucleotides.

There are two convenient operational modes in accordance with the present invention. In the first operational mode of the invention, the template nucleic acid is attached to a support. This can be either by immobilization of (1) an oligonucleotide primer or (2) a single-stranded or (3) double-stranded target nucleic acid molecule. Then, either (1) the target nucleic acid molecule is hybridized to the attached oligonucleotide primer, (2) an oligonucleotide primer is hybridized to the immobilized target nucleic acid molecule, to form a primed target nucleic acid molecule complex, or (3) a recognition site for the polymerase is created on the double stranded template (e.g., through interaction with accessory proteins, such as a primase). A nucleic acid polymerizing enzyme on the primed target nucleic acid molecule complex is provided in a position suitable to move along the target nucleic

acid molecule and extend the oligonucleotide primer at an active site. A plurality of labelled types of nucleotide analogs, which do not have a blocking substituent, are provided proximate to the active site, with each distinguishable type of nucleotide analog being complementary to a different nucleotide in the target nucleic acid sequence. The oligonucleotide primer is extended by using the nucleic acid polymerizing enzyme to add a nucleotide analog to the oligonucleotide primer at the active site, where the nucleotide analog being added is complementary to the nucleotide of the target nucleic acid at the active site. The nucleotide analog added to the oligonucleotide primer as a result of the extending step is identified. If necessary, the labeled nucleotide analog, which is added to the oligonucleotide primer, is treated before many further nucleotide analogs are incorporated into the oligonucleotide primer to insure that the nucleotide analog added to the oligonucleotide primer does not prevent detection of nucleotide analogs in subsequent polymerization and identifying steps. The steps of providing labelled nucleotide analogs, extending the oligonucleotide primer, identifying the added nucleotide analog, and treating the nucleotide analog are repeated so that the oligonucleotide primer is further extended and the sequence of the target nucleic acid is determined.

Alternatively, the above-described procedure can be carried out by first attaching the nucleic acid polymerizing enzyme to a support in a position suitable for the target nucleic acid molecule complex to move relative to the nucleic acid polymerizing enzyme so that the primed nucleic acid molecular complex is extended at an active site. In this embodiment, a plurality of labelled nucleotide analogs complementary to the nucleotide of the target nucleic acid at the active site are added as the primed target nucleic acid complex moves along the nucleic acid polymerizing enzyme. The steps of providing nucleotide analogs, extending the primer, identifying the added nucleotide analog, and treating the nucleotide analog during or after incorporation are repeated, as described above, so that the oligonucleotide primer is further extended and the sequence of the target nucleic acid is determined.

Figures 1A-C show 3 alternative embodiments for sequencing in accordance with the present invention. In Figure 1A, a sequencing primer is attached to a support, e.g. by a biotin-streptavidin bond, with the primer hybridized to the target nucleic acid molecule and the nucleic acid polymerizing enzyme attached to the

hybridized nucleic acid molecule at the active site where nucleotide analogs are being added to the sequencing primer. In Figure 1B, the target nucleic acid molecule is attached to a support, with a sequencing primer hybridized to the template nucleic acid molecule and the nucleic acid polymerizing enzyme attached to the hybridized nucleic molecule at the active site where nucleotide analogs are being added to the sequencing primer. The primer can be added before or during the providing of nucleotide analogs. In addition to these scenarios, a double stranded target nucleic acid molecule can be attached to a support, with the target nucleic acid molecule harboring a recognition site for binding of the nucleic acid polymerizing enzyme at an active site where nucleotide analogs are being added to the primer. For example, such a recognition site can be established with the help of an accessory protein, such as an RNA polymerase or a helicase/primase, which will synthesize a short primer at specific sites on the target nucleic acid and thus provide a starting site for the nucleic acid polymerizing enzyme. See Richardson "Bacteriophage T7: Minimal Requirements for the Replication of a Duplex DNA Molecule," Cell 33: 315-317 (1983), which is hereby incorporated by reference. In Figure 1C, the nucleic acid polymerizing enzyme is attached to a support, with the primed target nucleic acid molecule binding at the active site where nucleotide analogs are being added to the sequencing primer. As in the previous description, the nucleic acid polymerizing enzyme can likewise be attached to a support, but with the target nucleic acid molecule being double-stranded nucleic acid, harboring a recognition site for binding of the nucleic acid polymerizing enzyme at an active site where nucleotide analogs are being added to the growing nucleic acid strand. Although Figures 1A-C show only *one* sequencing reaction being carried out on the support, it is possible to conduct an array of several such reactions at different sites on a single support. In this alternative embodiment, each sequencing primer, target nucleic acid, or nucleic acid polymerizing enzyme to be immobilized on this solid support is spotted on that surface by microcontact printing or stamping, e.g., as is used for microarray technology of DNA chips, or by forming an array of binding sites by treating the surface of the solid support. It is also conceivable to combine the embodiments outlined in Figure 1 and immobilize both the target nucleic acid molecule and the nucleic acid polymerizing enzyme proximate to each other.

through lens 6. Dichroic beam splitter 8 allows passage of the emitted radiation to detector 12 which identifies the type of emission. The detected emission information is then directed to computer 14 where the nucleotide base corresponding to the emission is identified and its identity stored. After multiple cycles of this procedure, the computer will be able to generate as output the sequence of the target nucleic acid molecule.

Figures 10A-B show systems for supplying reagents to a nanofabricated confinement system in accordance with the present invention. In Figure 10A, the reagents, which include dATP, dCTP, dGTP, dUTP, the nucleic acid source, and buffer are held in separate reservoirs and connected through separate conduits to manifold 200 where the reagents are mixed together before entering nanochannel 202. The components of this system upstream and downstream of nanochannel 202 can be combined as a microstructure. In the process of passing rapidly through nanochannel 202, the reagents move rapidly through reaction zone 204 where the sequencing procedure of the present invention is carried out. From nanochannel 202, the residual reagents R pass through outlet 206. The system of Figure 10B is generally similar to that of Figure 10A, but the former system is on a single chip with pads to connect the system to fluid reservoirs. In particular, the reservoir for each of the reagents is coupled to the chip 208 via inlet pads 210a-f, while the outlet for discharged reagents is connected to pad 212.

Nanofabricated channels of 75 nm width and 60 nm height have been manufactured with excellent optical transparency and used for DNA flow control. See Turner et al., "Solid-State Artificial Gel for DNA Electrophoresis with an Integrated Top Layer," Proceedings of SPIE: Micro- and Nano-Fabricated Structures and Devices for Biomedical Environmental Applications 3258:114-121 (1998), which is hereby incorporated by reference. By placing the nucleic acid synthesis complex into a channel of depth z = 25 nm, minimizing the x-dimension of the focused laser beam to ca. 300 nm, and fixing the y-dimension by the channel width at 100 nm, the effective volume of observation can be reduced to 7.5×10^{-4} μm^3, corresponding to 0.75 attoliters. Here, the concentration for only one substrate molecule to be present in the excitation volume amounts to 2 μM, a substrate concentration well within the range of rapid and efficient nucleic acid polymerization. Moreover, since there are

four different nucleotide analogs, each to be distinguished, the effective substrate concentration for the polymerase is four times higher. If a smaller effective volume of observation is required, the y-dimension in the flow direction can be reduced to about 100 nm by illumination with the interference pattern of two objectives at about 90° axial angles as in theta microscopy. See Stelzer et al., "A New Tool for the Observation of Embryos and Other Large Specimens: Confocal Theta Fluorescence Microscopy," J. Microscopy 179:1-10 (1995), which is hereby incorporated by reference.

To excite the labels, activating energy is focused proximate to the active site of polymerase extension (i.e. where the polymerase is located). To the extent this active site moves during extension (e.g., as a result of movement by the polymerase), the focus of the activating energy is also moved.

A necessary consideration is the choice between one-photon and multiphoton excitation of fluorescence. Multiphoton excitation provides some powerful advantages, but it is more complex and more expensive to implement. Multiphoton excitation fluorescence utilizing simultaneous absorption of two or more photons from bright, femtosecond infrared pulses generated by ultrafast solid state mode locked lasers provides the most promising approach. See Denk et al., "2-Photon Laser Scanning Fluorescence Microscopy," Science 248:73-76 (1990), which is hereby incorporated by reference. Sensitivity to single molecule fluorescence is routinely obtained and is temporally resolvable to the microsecond level with fluorescence lifetimes measurable with reasonable accuracy for single molecules. See Mertz et al., "Single-Molecule Detection by Two-Photon-Excited Fluorescence," Optics Lett. 20:2532-2534 (1995) and Eggeling et al., "Monitoring Conformational Dynamics of a Single Molecule by Selective Fluorescence Spectroscopy," Proc. Natl. Acad. Sci. USA 95:1556-1561 (1998), which are hereby incorporated by reference.

The ideal fluorescent signal for single molecule sequencing consists of time resolved bursts of distinguishable fluorescence as each nucleotide is bound. Thus, in the ideal situation, a time-resolved train of color resolved fluorescent bursts could be obtained if nucleotides were bound at distinguishable intervals as described in Figure 3. Full resolution of the time sequence of events therefore offers the best background reduction and reliable possibility for nucleotide recognition. Since with

the currently available polymerases, labelled nucleotides are most likely added no faster than at 1 millisecond intervals, it should be possible that all of the detected fluorescence photons from each labelled nucleotide can be accumulated and removed before the next fluorescent nucleotide is bound. This ideal burst-gap-burst sequence is realized although actually every molecular kinetic step of polymerization involves the stochastic Poisson process. For a single Poisson process, the most probable time delay between events is zero although the average delay would be larger than zero. However, the process of incorporation of a single dNTP into DNA by DNA polymerase is a sequential multistep process of at least 5 different events. See Patel et al., "Pre-Steady-State Kinetic Analysis of Processive DNA Replication Including Complete Characterization of an Exonuclease-Deficient Mutant," Biochemistry 30: 511-525 (1991). The sequential summation of these steps will result in a most likely time delay larger than zero. Therefore, the photon bursts are not very likely to overlap.

For conventional fluorophores, about 10^5 photons per fluorophore will be emitted before photobleaching. Detection of (at most) 1% of the emission yields about 10^3 photons for a relative noise uncertainty of 3%. Background due to free nucleotides is reduced to a nearly negligible level by the schemes discussed above, e.g., by limiting the size of the focal volume to contain only about one free labelled nucleotide, with very short dwell times.

The expected detection level is about 10^3 photons from each labelled nucleotide, in about 10^{-3}s. This is an acceptable counting rate, ~ 10^6 Hz, and an acceptable fluorophore excitation rate at about one tenth of singlet excited state saturation. This fluorescence excitation creates a detected burst of ~10^3 photons in about 1 ms at the characteristic wavelength for each labelled nucleotide, leaving, on average, a gap of about 1 ms before the next nucleotide is added, well within the average time intervals between nucleotide addition at probably more than one millisecond. Possible burst overlaps can be analyzed and resolved by the analytical treatment of continuous measurements of data in time coherent sequences in (at best) 4 channels for most accurate sequencing results. With the photon statistics available in the experimental design and recently developed coupled multichannel analyzers and operational software, error rates can be made acceptable with 4 labelled

nucleotides or with the strategies involving a smaller number of labels as outlined above.

Spectral resolution of four fluorophores identifying the nucleotides can be achieved with two-photon excitation by infrared pulses. All 4 fluorophores can be simultaneously excited due to the wide excitation bands usually characteristic of two-photon excitation. See Xu et al., "Multiphoton Excitation Cross-Sections of Molecular Fluorophores," Bioimaging 4:198-207 (1996), which is hereby incorporated by reference. Alternatively, multiple excitation sources can be used in combination or by fast switching to illuminate the sequencing complex if necessary. Spectral separation is accomplished with conventional interference filters but emission spectra may overlap, complicating the time correlation analysis and perhaps requiring cross correlation of the 4 color channels for correction. If compatibility of fluorophores with the nucleic acid polymerizing enzyme limits the applicability of suitable dye sets, a combination of techniques can be applied to distinguish the labels.

Another potential way to distinguish incorporation of a nucleotide into the growing nucleic acid strand consists of measuring changes in fluorescence lifetime. Fluorescence lifetime of an oligonucleotide pyrene probe has been observed to vary in a sequence-dependent manner upon DNA attachment. See Dapprich J, "Fluoreszenzdetection Molekularer Evolution (Fluorescence Detection of Molecular Evolution)," Dissertation, Georg-August-Univ., Goettingen, Germany (1994), which is hereby incorporated by reference. Photophysical interactions between the fluorophore and the base result in characteristic fluorescence decay times, and can also be used to differentiate the bases, as discussed above. Lifetime determination and discrimination on the single molecule level has recently been demonstrated so that discrimination between bases being incorporated and freely diffusing nucleotides could be carried out by fluorescence lifetime measurements. See Eggeling et al., "Monitoring Conformational Dynamics of a Single Molecule by Selective Fluorescence Spectroscopy," Proc. Natl. Acad. Sci. USA 95:1556-1561 (1998), which is hereby incorporated by reference.

Time correlated measurements in four fluorescence wavelength channels can be used effectively in carrying out the process of the present invention. Overlap of emission spectra may allow signals from one fluorophore to enter several

channels but the relative count rate and timing identifies the label correctly. Simultaneous signals from an incorporated labelled nucleotide and a free label are distinguishable by the time duration and magnitude of the bursts, which are limited for the free label. Label ambiguity can be further reduced by utilization of fluorescence decay time measurements which can be realized with the available 0.1 ns resolution of time delays for fluorescence photon emission after each femtosecond laser excitation pulse. The fluorescence photon emission and photobleaching processes themselves are also stochastic processes but involve sufficiently disparate quantum efficiencies that error rates should be negligible.

In rejecting background from the freely diffusing or flowing labelled nucleotides, the very short dwell time of any individual free nucleotide in the focal volume is advantageously used. The characteristic diffusion time for a free nucleotide analog across the open dimension of the focal volume (in the worst case of non-interferometric far-field illumination) will be $\tau_D \sim y^2/4D \sim 2 \times 10^{-5}$ sec, with y being the focal volume dimension and D the diffusion coefficient. An iontophoretic flow velocity of 1cm/s is sufficient to keep its short bursts of fluorescence to less than 10^{-5} sec and reduce the photon numbers by an order of magnitude. This will assure discrimination against free nucleotides and identify the time series of bursts representing the nucleic acid sequence, provided the nucleotide analog concentrations are appropriately low as discussed. Magde et al., "Thermodynamic Fluctuations in a Reacting System – Measurement by Fluorescence Correlation Spectroscopy," Phys. Rev. Lett. 29:705-708 (1972) and Maiti et al., "Measuring Serotonin Distribution in Live Cells with Three-Photon Excitation," Science 275:530-532 (1997), which are hereby incorporated by reference. Discrimination can be improved by utilizing volume confinement techniques or time-gated detection, as discussed above.

Detection of fluorescence resonance energy transfer (FRET) from a donor fluorophore (e.g., a donor attached to the polymerase) to adjacent nucleotide analog acceptors that are incorporated into the growing nucleic acid strand suggests a further elegant possibility of lowering background from incorporated nucleotides. FRET only reaches very short distances including about 20 nucleotides and decays at the reciprocal sixth power of distance. The excited donor molecule transfers its energy only to nearby acceptor fluorophores, which emit the spectrally resolved

acceptor fluorescence of each labelled nucleotide as it is added. Already incorporated nucleotides farther away from the donor would not contribute to the fluorescent signal since distance and orientation constraints of energy transfer reduce the effective range of observation to less than 60 Å, thereby effectively eliminating background fluorescence from unincorporated nucleotides. Without photobleaching, the method requires high sensitivity since repeat nucleotides leave the range of FRET at the same rate that new nucleotides are added, possibly creating sequence recognition ambiguity. Photobleaching or photochemical cleavage, or their combination as discussed above could resolve the problem. Photobleaching of the donor molecules using FRET can be avoided if it is the template nucleic acid that is attached and the donor bearing nucleic acid polymerizing enzyme is periodically replaced.

A final important consideration for the success of the present invention concerns the stability of the protein/nucleic acid complex in activating radiation, such as tightly focussed laser beams. It is not expected that the enzyme is affected by the excitation illumination, because wavelengths are chosen at which proteins do not absorb, the stability of the polymerase in the laser beam should be sufficiently high to allow for accurate sequencing runs over long read lengths. Previous investigations exposing enzymes to strong laser light have examined photodamage and loss of function. Immobilized RNA polymerase/DNA complexes showed inactivation times of 82 ± 58 s for 1047 nm Nd:Y laser light of 82 to 99 mW laser power focused at the protein, corresponding to intensities of approximately 10^8 W/cm^2. Other studies on the actomyosin or kinesin systems indicated similar stability. Both DNA and biotin-avidin linkages have been shown to be photostable in optical traps. See Yin et al., "Transcription Against an Applied Force," Science 270: 1653-1657 (1995), Svoboda et al. "Direct Observation of Kinesin Stepping by Optical Trapping Interferometry," Nature 365: 721-727 (1993), and Molloy et al., "Movement and Force Produced by a Single Myosin Head" Nature 378: 209-212 (1995), which are hereby incorporated by reference. For fluorescence detection of nucleotide analogs according to the present invention, laser powers (intensities) typical of FCS measurements are expected, on the order of 0.1 mW (10^5 W/cm^2) for one-photon and 1 mW (10^6-10^7 W/cm^2) for two-photon excitation, thereby being significantly lower than in the case of optical tweezers described above. Enzyme stability should therefore be higher, moreover,

with the rapid speed of sequencing proposed by this method (e.g., 100 bp/s), even 80 s are sufficient to determine the sequence of 8 kb nucleic acid.

Although the invention has been described in detail for the purposes of illustration, it is understood that such detail is solely for that purpose, and variations can be made therein by those skilled in the art without departing from the spirit and scope of the invention which is defined by the following claims.

WHAT IS CLAIMED:

1. A method of sequencing a target nucleic acid molecule having a plurality of nucleotide bases comprising:

providing a complex of a nucleic acid polymerizing enzyme and the target nucleic acid molecule oriented with respect to each other in a position suitable to add a nucleotide analog at an active site complementary to the target nucleic acid;

providing a plurality of types of nucleotide analogs proximate to the active site, wherein each type of nucleotide analog is complementary to a different nucleotide in the target nucleic acid sequence;

polymerizing a nucleotide analog at an active site, wherein the nucleotide analog being added is complementary to the nucleotide of the target nucleic acid, leaving the added nucleotide analog ready for subsequent addition of nucleotide analogs;

identifying the nucleotide analog added at the active site as a result of said polymerizing; and

repeating said providing a plurality of types of nucleotide analogs, said polymerizing, and said identifying so that the sequence of the target nucleic acid is determined.

2. A method according to claim 1, wherein the nucleic acid polymerizing enzyme is selected from the group consisting of a DNA polymerase, an RNA polymerase, reverse transcriptase, and mixtures thereof.

3. A method according to claim 1, wherein the nucleic acid polymerizing enzyme is a thermostable polymerase.

4. A method according to claim 1, wherein the nucleic acid polymerizing enzyme is a thermodegradable polymerase.

5. A method according to claim 1, wherein the target nucleic acid molecule is selected from the group consisting of double-stranded DNA, single-stranded DNA, single stranded DNA hairpins, DNA/RNA hybrids, RNA with a recognition site for binding of the polymerase, and RNA hairpins.

6. A method according to claim 1, wherein the nucleic acid polymerizing enzyme is bound to the target nucleic acid molecule complex at an origin of replication, a nick or gap in a double-stranded target nucleic acid, a secondary structure in a single-stranded target nucleic acid, a binding site created by an accessory protein, or a primed single stranded nucleic acid.

7. A method according to claim 1, wherein the nucleic acid polymerizing enzyme is provided with one or more accessory proteins to modify its activity.

8. A method according to claim 7, wherein the accessory protein is selected from the group consisting of a single-stranded binding protein, a primase, and helicase.

9. A method according to claim 1, wherein the nucleic acid polymerizing enzyme is processive.

10. A method according to claim 1, wherein the nucleic acid polymerizing enzyme is non-processive.

11. A method according to claim 1, wherein the nucleotide analogs are selected from the group consisting of a ribonucleotide, a deoxyribonucleotide, a modified ribonucleotide, a modified deoxyribonucleotide, a peptide nucleotide, a modified peptide nucleotide, and a modified phosphate-sugar backbone nucleotide.

12. A method according to claim 1 further comprising:

hybridizing an oligonucleotide primer to the target nucleic acid molecule prior to or during said providing a plurality of nucleotide analogs.

13. A method according to claim 12, wherein the oligonucleotide primer comprises nucleotides selected from the group consisting of ribonucleotides, deoxyribonucleotides, modified ribonucleotides, modified deoxyribonucleotides, peptide nucleic acids, modified peptide nucleic acids, and modified phosphate-sugar backbone nucleotides.

14. A method according to claim 1, wherein the nucleotide analogs are provided with a label.

15. A method according to claim 14, wherein the label is selected from the group consisting of chromophores, fluorescent moieties, enzymes, antigens, heavy metals, magnetic probes, dyes, phosphorescent groups, radioactive materials, chemiluminescent moieties, scattering or fluorescent nanoparticles, Raman signal generating moieties, and electrochemical detection moieties.

16. A method according to claim 14, wherein the label is attached to the nucleotide analog at its base, sugar moiety, alpha phosphate, beta phosphate, or gamma phosphate.

17. A method according to claim 14, wherein the label is attached to the nucleotide analog with a linker.

18. A method according to claim 14, wherein the label is attached to the nucleotide analog without a linker.

19. A method according to claim 14 further comprising:

removing the label from the nucleotide analog during or after said identifying and before said polymerizing many further nucleotide analogs at the active site.

20. A method according to claim 19, wherein said removing is carried out by bleaching the label.

21. A method according to claim 20, wherein said bleaching is carried out by photobleaching with radiation which is adjusted to induce and control label removal.

22. A method according to claim 19, wherein said removing is carried out by cleaving the label from the nucleotide analog.

23. A method according to claim 22, wherein beta- or gamma-labeled nucleotide analogs are enzymatically cleaved.

24. A method according to claim 14, wherein each of the plurality of types of nucleotide analogs have different labels which are distinguished from one another during said identifying.

25. A method according to claim 14, wherein three or less of the plurality of types of nucleotide analogs have a different label.

26. A method according to claim 14, wherein the different types of nucleotide analogs have the same label but are distinguished by different properties due to the presence of base fluorophores, quenched fluorophores, or fluorogenic nucleotide analogs.

27. A method according to claim 1, wherein the nucleic acid polymerizing enzyme carries a label and said identifying is carried out by detecting interaction between the label and the nucleotide analog.

28. A method according to claim 27, wherein the label is a fluorescence resonance energy transfer donor or acceptor.

29. A method according to claim 1, wherein said identifying is carried out by non-optical procedures.

30. A method according to claim 1, wherein said identifying is carried out by optical procedures selected from the group consisting of far-field microspectroscopy, near-field microspectroscopy, evanescent wave or wave guided illumination, nanostructure enhancement, and combinations thereof.

31. A method according to claim 1, wherein said identifying is carried out by utilizing single and/or multiphoton excitation, fluorescence resonance energy transfer, or photoconversion.

32. A method according to claim 1, wherein said identifying is achieved by spectral wavelength discrimination, measurement and separation of fluorescence lifetimes, fluorophore identification and/or background suppression.

33. A method according to claim 32, wherein fluorophore identification and/or background suppression utilizes fast switching between excitation modes and illumination sources, and combinations thereof.

34. A method according to claim 1, wherein said providing a complex comprises:

positioning either (1) an oligonucleotide primer or (2) the target nucleic acid molecule on a support;

hybridizing either (1) the target nucleic acid molecule to the positioned oligonucleotide primer or (2) an oligonucleotide primer to the positioned target nucleic acid molecule, to form a primed target nucleic acid molecule complex; and

providing the nucleic acid polymerizing enzyme on the primed target nucleic acid molecule complex in a position suitable to move along the target nucleic acid molecule and extend the oligonucleotide primer at an active site.

35. A method according to claim 34, wherein said hybridizing is carried out by additionally binding the end of the target nucleic acid molecule opposite to that bound to the oligonucleotide primer to a second oligonucleotide primer positioned on the support.

36. A method according to claim 34, wherein the support and either the oligonucleotide primer or the target nucleic acid molecule are bound reversibly or irreversibly with corresponding components of a covalent or non-covalent binding pair selected from the group consisting of an antigen-antibody binding pair, a streptavidin-biotin binding pair, photoactivated coupling molecules, and a pair of complementary nucleic acids.

37. A method according to claim 34, where the oligonucleotide primer is positioned on the support and the target nucleic acid molecule is hybridized to the positioned oligonucleotide primer.

38. A method according to claim 34, wherein the target nucleic acid molecule is positioned on the support and the oligonucleotide primer is hybridized to the positioned target nucleic acid molecule.

39. A method according to claim 1, wherein said providing a complex comprises:

positioning, on a support, a double stranded nucleic acid molecule comprising the target nucleic acid and having a recognition site proximate the active site, and

providing the nucleic acid polymerizing enzyme on the target nucleic acid molecule in a position suitable to move along the target nucleic acid molecule.

40. A method according to claim 1, wherein said providing a complex comprises:

positioning a nucleic acid polymerizing enzyme on a support in a position suitable for the target nucleic acid complex to move relative to the nucleic acid polymerizing enzyme.

41. A method according to claim 40, wherein the support and the nucleic acid polymerizing enzyme are bound reversibly or irreversibly with corresponding components of a covalent or non-covalent binding pair selected from the group consisting of an antigen-antibody binding pair, a streptavidin-biotin binding pair, photoactivated coupling molecules, and a pair of complementary nucleic acids.

42. A method according to claim 1, wherein the nucleic acid polymerizing enzyme or the target nucleic acid is positioned on an adjustable support.

43. A method according to claim 1, wherein the nucleic acid polymerizing enzyme or the target nucleic acid is positioned in a gel with pores.

44. A method according to claim 1, wherein the target nucleic acid and the nucleic acid polymerizing enzyme are positioned on a solid support proximate to each other.

45. A method according to claim 1, wherein said identifying is carried out by reducing background noise resulting from free nucleotide analogs.

46. A method according to claim 45, wherein said identifying comprises:

directing activating radiation to a region substantially corresponding to the active site and

detecting the nucleotide analog polymerized at the active site.

47. A method according to claim 45, wherein said identifying distinguishes nucleotide analogs polymerized at the active site from free nucleotide analogs.

48. A method according to claim 45, wherein said identifying is carried out in a confined region proximate to the active site.

49. A method according to claim 48, wherein said identifying is carried out in a nanostructure.

50. A method according to claim 49, wherein the nanostructure is a punctuate, acicular, or resonant nanostructure which enhances said detecting.

51. A method according to claim 48, wherein nucleotide analogs that are not polymerized at the active site move rapidly through a microstructure to and from the confined region.

52. A method according to claim 51, wherein the microstructure comprises:

a plurality of channels to direct different nucleotide analogs to the confined region and

a discharge channel to permit materials to be removed from the confined region, and the nanostructure comprises:

a housing defining the confined region and constructed to facilitate said identifying.

53. A method according to claim 45, wherein said identifying is carried out by electromagnetic field enhancement with electromagnetic radiation being enhanced proximate to an object with a small radius of curvature adjacent to the active site.

54. A method according to claim 45, wherein said identifying is carried out by near-field illumination of cavities in which the primed target nucleic acid molecule is positioned.

55. A method according to claim 45, wherein said identifying is carried out with optical fibers proximate to the complex.

56. A method according to claim 45, wherein said identifying and said reducing background is carried out by time gated delay of photon detection.

57. A method according to claim 1, wherein said method is carried out by sequencing different target nucleic acid molecules at a plurality of different locations on an array.

58. A method according to claim 1, wherein said method is carried out by simultaneously or sequentially sequencing the same target nucleic acid and combining output from such sequencing.

59. An apparatus suitable for sequencing a target nucleic acid molecule comprising:

a support;

a nucleic acid polymerizing enzyme or oligonucleotide primer suitable to bind to a target nucleic acid molecule, wherein said nucleic acid polymerizing enzyme or said oligonucleotide primer is positioned on said support; and

a microstructure defining a confined region containing said support and said nucleic acid polymerizing enzyme or said oligonucleotide primer and configured to permit labeled nucleotide analogs that are not positioned on the support to move rapidly through the confined region.

60. An apparatus according to claim 59, wherein the microstructure comprises:

a plurality of channels to direct different types of nucleotide analogs to the confined region and

a discharge channel to permit materials to be removed from the confined region and a nanostructure constructed to facilitate identification of nucleotide analogs positioned on the support.

61. An apparatus suitable for sequencing a target nucleic acid molecule comprising:

a support;

a nucleic acid polymerizing enzyme or oligonucleotide primer suitable to hybridize to a target nucleic acid molecule, wherein said nucleic acid polymerizing enzyme or said oligonucleotide primer is positioned on said support;

a housing defining a confined region containing said support and said nucleic acid polymerizing enzyme or said oligonucleotide primer and constructed to facilitate identification of labeled nucleotide analogs positioned on the support; and

optical waveguides proximate to the confined region to focus activating radiation on the confined region and to collect radiation from the confined region.

Figure 1

A
ssDNA template
DNA polymerase
primer
reptavidin
biotin
support surface
B
DNA polymerase
ssDNA template
primer
streptavidin
biotin
support surface
C
DNA polymerase
ssDNA template
primer
support surface

Figure 2

Figure 3

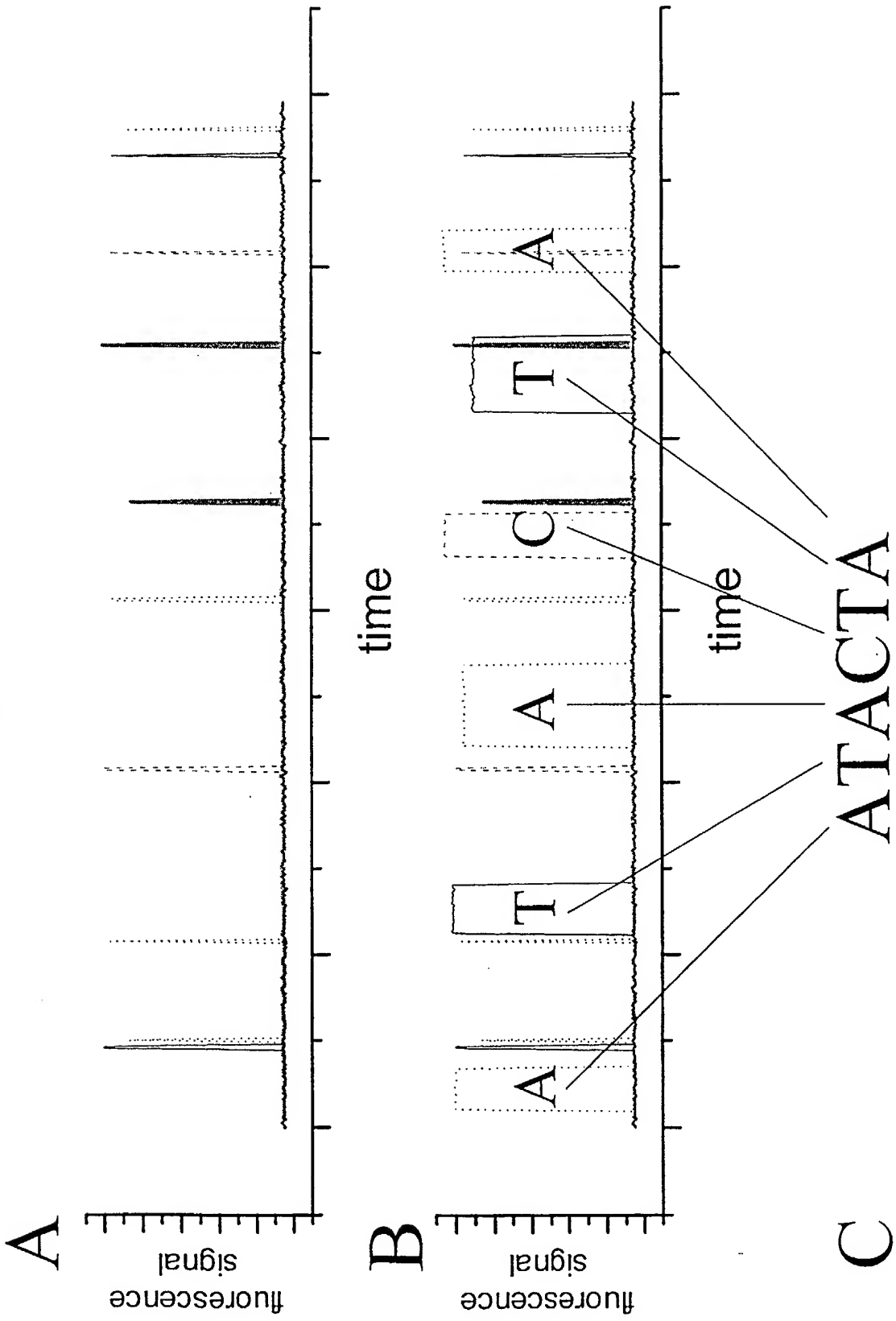

A
fluorescence signal
time
B
fluorescence signal
time
A
T
A
C
T
A
C
ATACTA

Figure 4

Figure 5

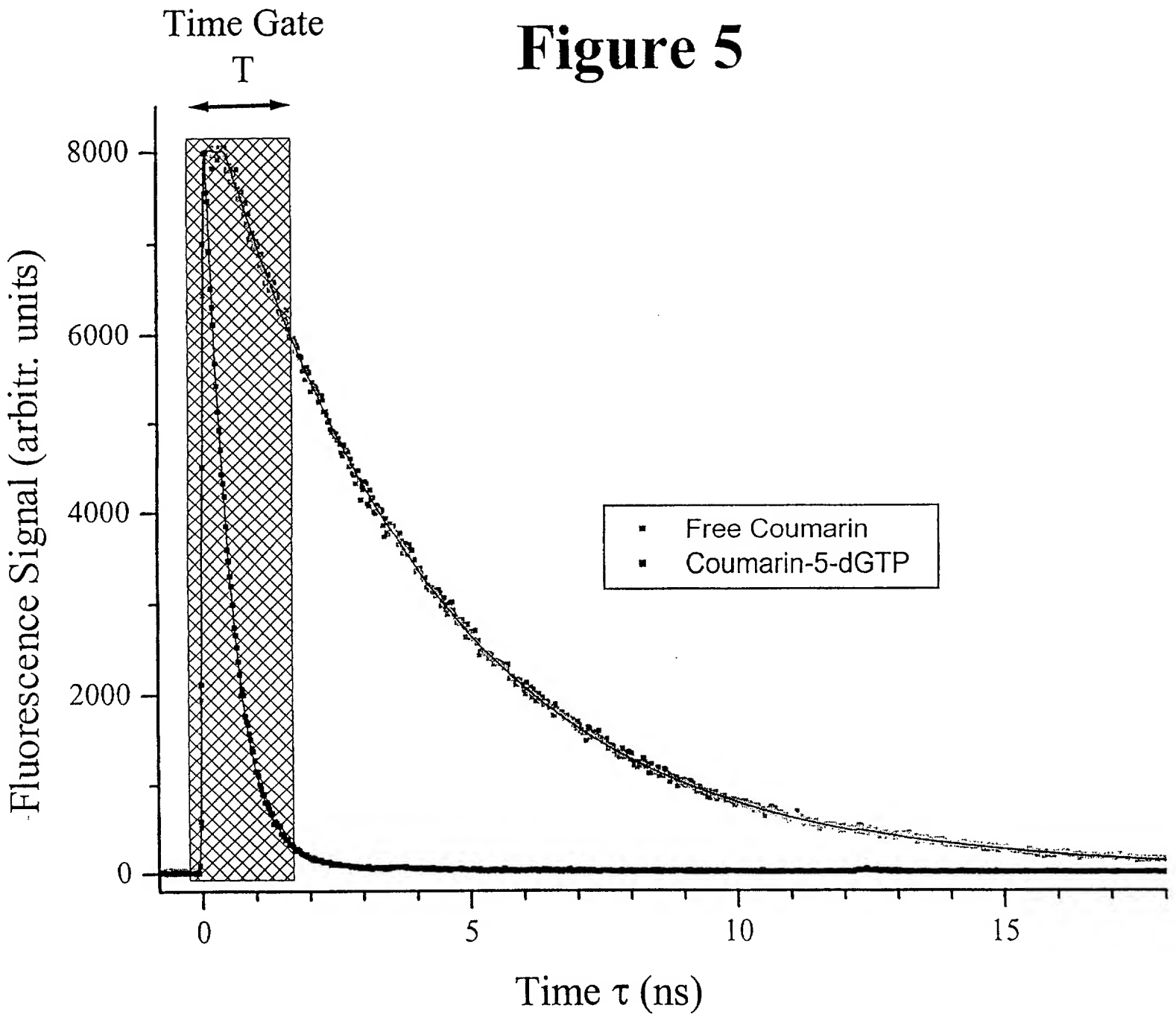

Time Gate
T
Fluorescence Signal (arbitr. units)
8000
6000
4000
2000
0
0
5
10
15
Time τ (ns)
Free Coumarin
Coumarin-5-dGTP

Figure 6

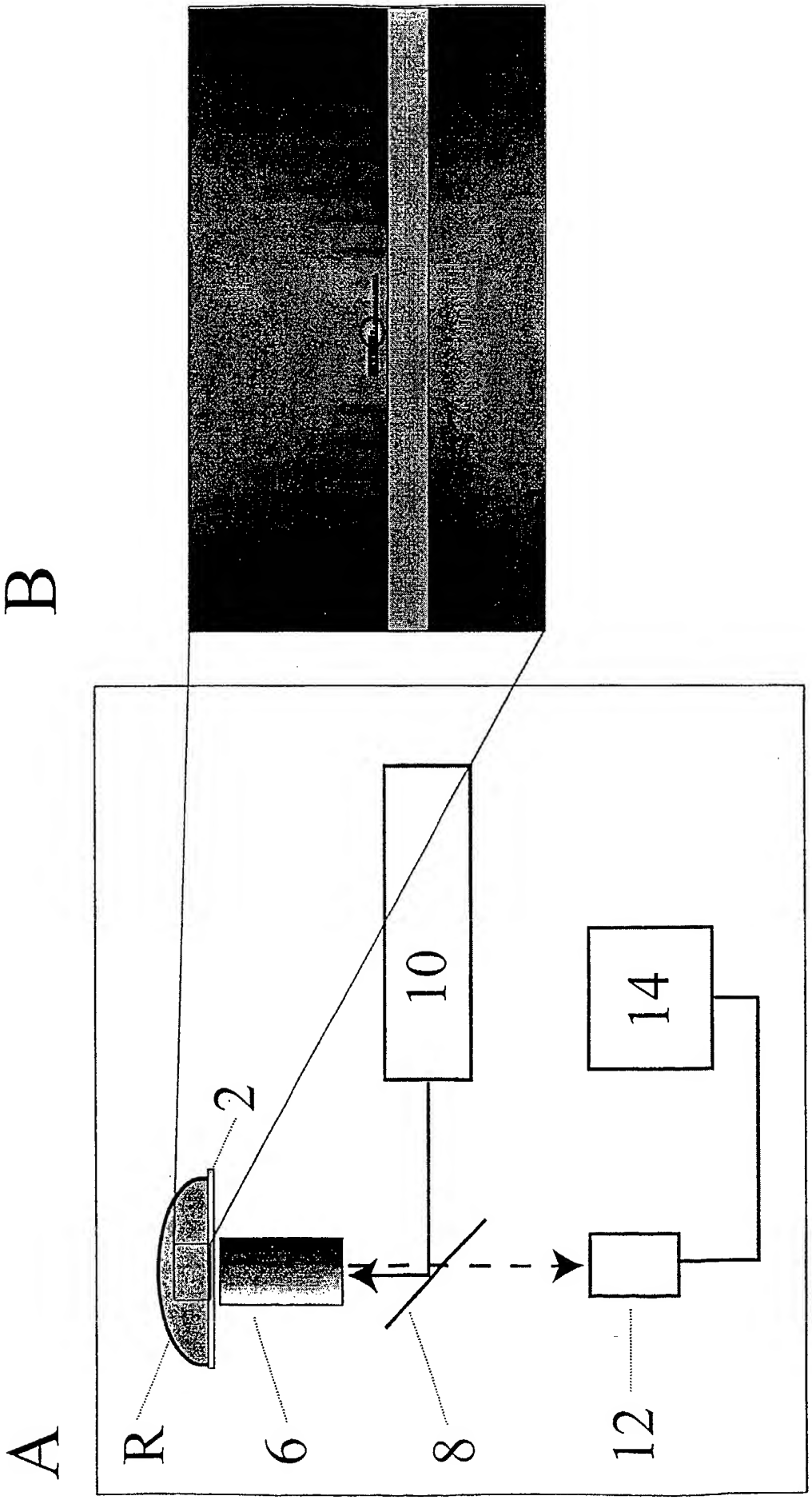


Figure 7

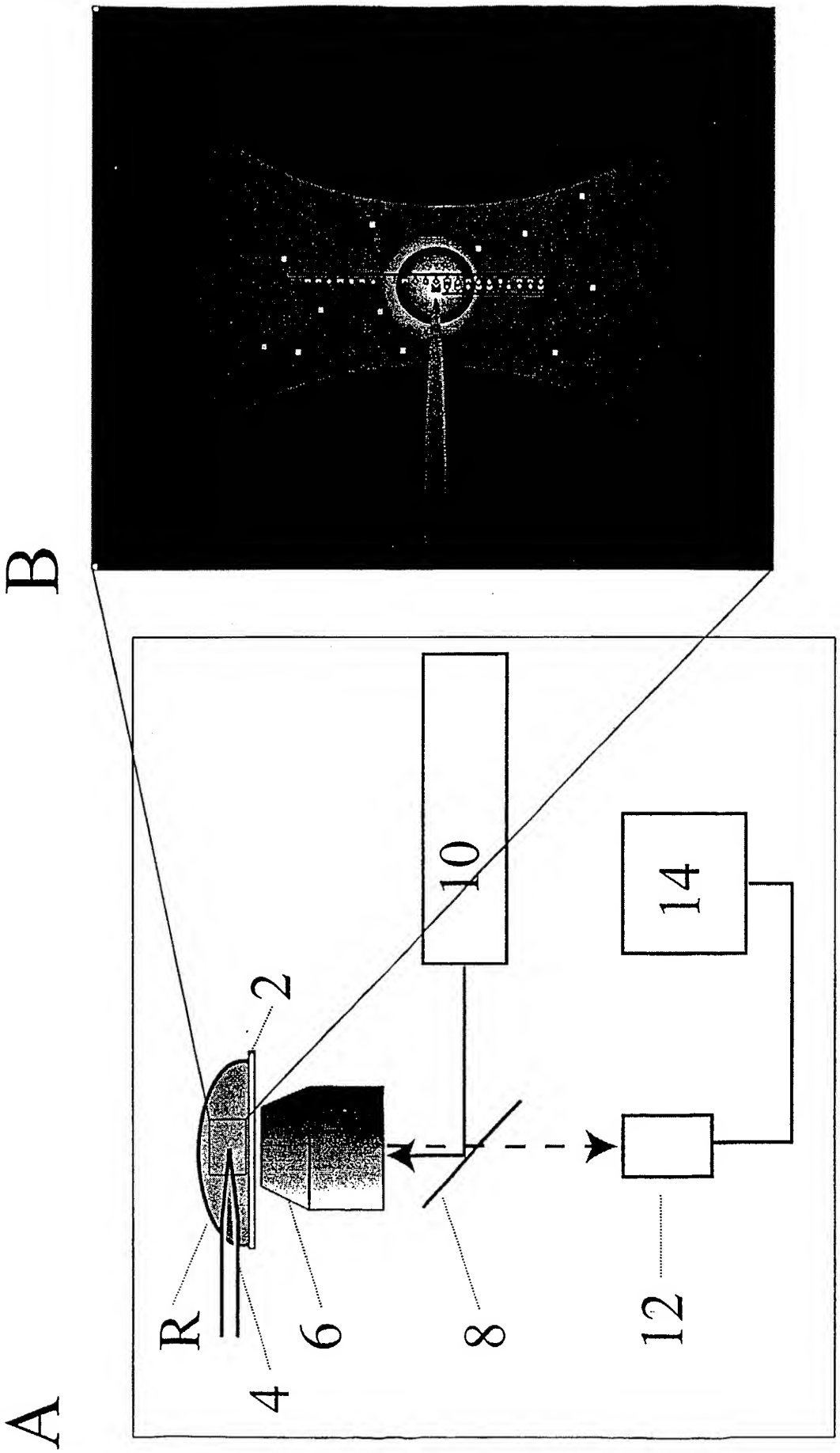


Figure 8

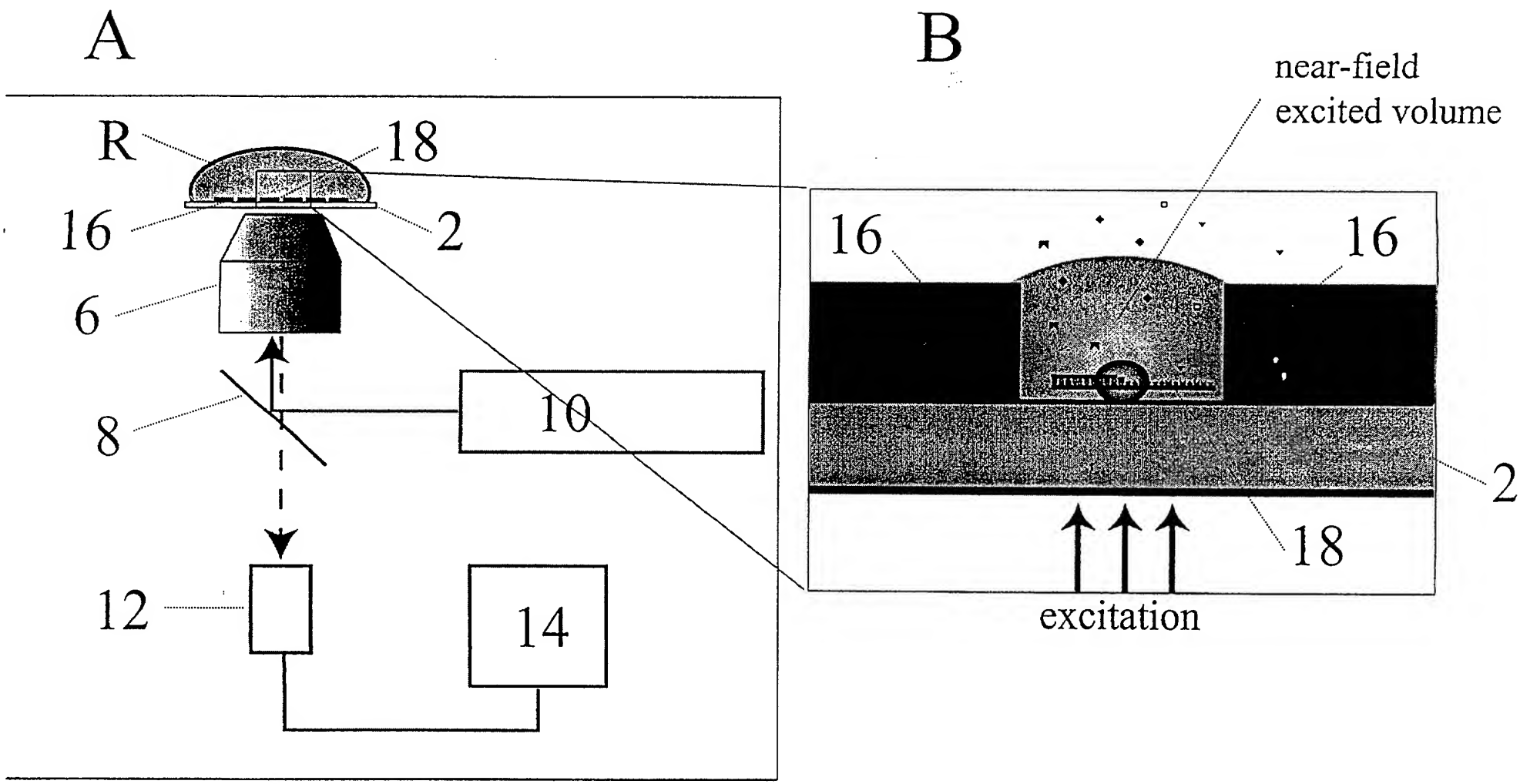

A
R
18
16
2
6
8
10
12
14
B
near-field
excited volume
16
16
2
18
excitation

Figure 9

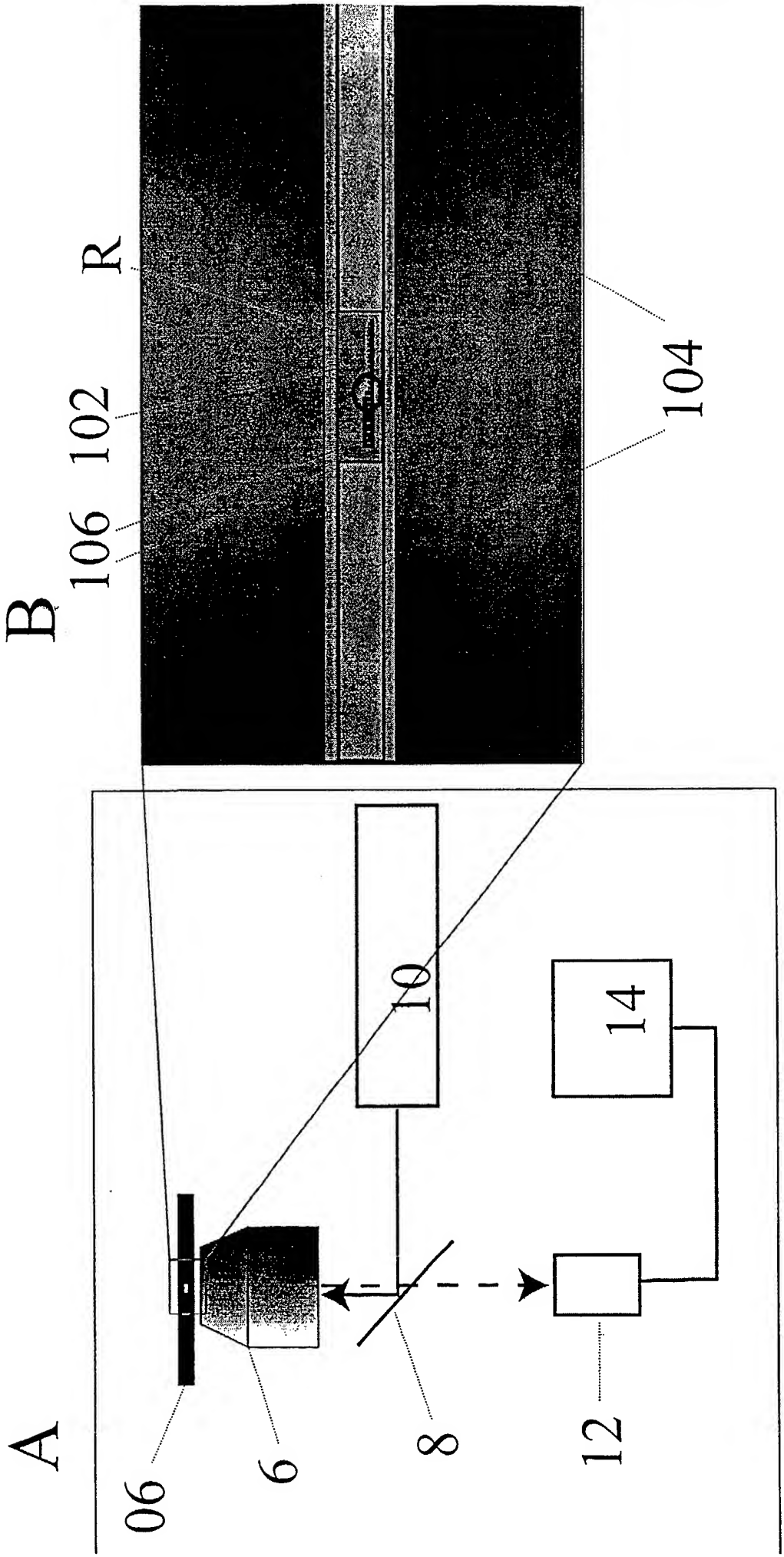
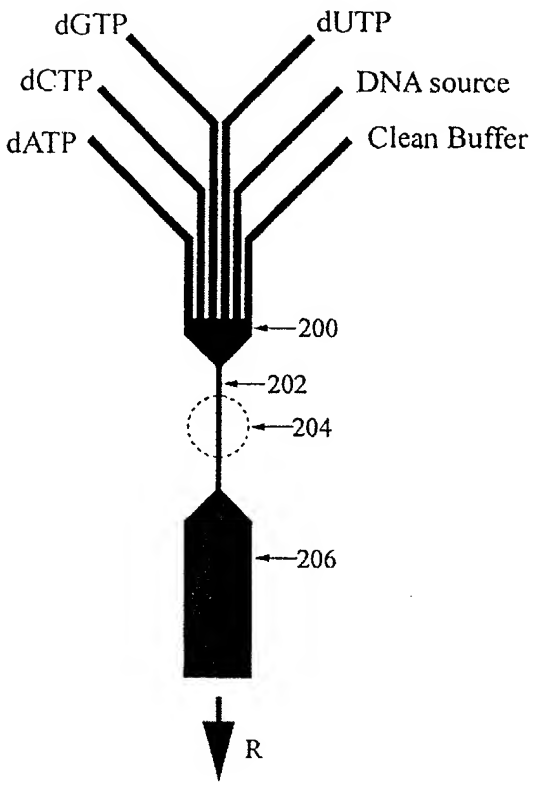
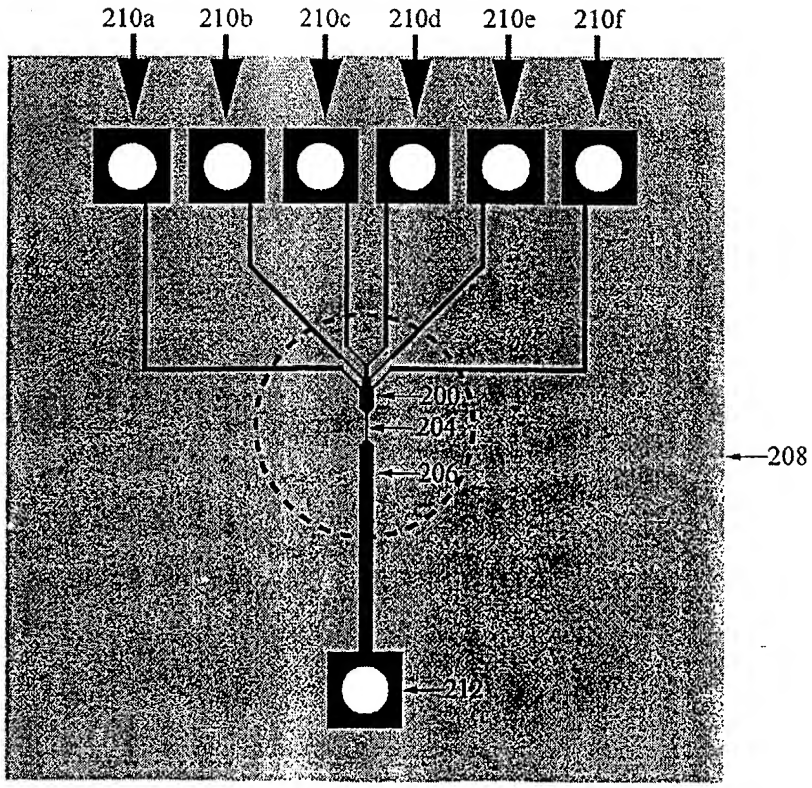


Figure 10

A.


dGTP
dUTP
dCTP
DNA source
dATP
Clean Buffer
200
202
204
206
R

B.


210a
210b
210c
210d
210e
210f
200
204
206
208
212

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/13677

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) :C12P 19/34; C12Q 1/68; C12M 1/34
US CL : 435/6, 91.1, 287.2
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 91.1, 183, 287.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,470,710 A (WEISS et al.) 28 November 1995, see entire document.	1-18, 24-58
Y	US 5,547,835 A (KOSTER) 20 April 1996, see entire document.	1-19, 24-58
A,P	US 5,961,923 A (NOVA et al.) 05 October 1999, see entire document.	1-61
A,P	US 6,027,890 A (NESS et al.) 22 February 2000, see entire document.	1-61
A,P	US 6,048,690 A (HELLER et al.) 11 April 2000, see entire document.	1-61

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
11 JULY 2000

Date of mailing of the international search report
01 AUG 2000

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
BRADLEY L. SISSON
Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/13677

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

EAST

Search Terms: nucleic acid, multiples, fluorescent, nucleotide, analog?, bleaching